

On the Inverse Power Flow Problem

Y. Yuan, *Member, IEEE*, O. Ardakanian, *Member, IEEE*, S. H.

Low, *Fellow, IEEE*, and C. J. Tomlin, *Fellow, IEEE*

Abstract

This paper studies the inverse power flow problem which is to infer the nodal admittance matrix and the network topology from voltage and current magnitudes and phase angles measured at a number of buses. We show that the admittance matrix can be uniquely identified from a sequence of measurements of different steady states when the system is fully observable, and a reduced admittance matrix (from Kron reduction) can be determined when the system is not fully observable (i.e., it contains some hidden nodes). Furthermore, we propose efficient algorithms based on graph theory and optimization to uncover the original admittance matrix of radial and mesh systems with hidden nodes. As a case study, we explore the application of the proposed algorithms to detect and locate faults and other critical events. Simulations performed on a standard test system confirm that these algorithms provide accurate estimates of the admittance matrix even when the measurements are perturbed by additive white Gaussian noise.

Index Terms

Inverse Power Flow Problem, System Identification, Phasor Measurement Units, Fault Detection.

I. INTRODUCTION

The power industry has witnessed profound changes in recent years, which are mostly driven by the widespread adoption of distributed energy resources (DER), active participation of customers in emerging energy markets, and rapid deployment of measurement, communication, and control infrastructure resulting in an unprecedented level of visibility and controllability,

Ye Yuan is with School of Automation, Huazhong University of Science and Technology, Wuhan, China. Omid Ardakanian and Claire J. Tomlin are with Department of Electrical Engineering and Computer Sciences, University of California, Berkeley. Steven H. Low is with Department of Computer and Mathematical Sciences and Electrical Engineering, California Institute of Technology. For correspondence, contact slow@caltech.edu.

This work is supported in part by NSF under CPS:ActionWebs (CNS-0931843) and CPS:FORCES (CNS1239166), by NASA under grants NNX12AR18A and UCSCMCA-14-022 (UARC), by ONR under grants N00014-12-1-0609, N000141310341 (Embedded Humans MURI), and MIT_5710002646 (SMARTS MURI), and by AFOSR under grants UPenn-FA9550-10-1-0567 (CHASE MURI) and the SURE project.

especially for distribution grids. These changes offer ample opportunities to system operators to improve power system efficiency and stability, despite the increased amount of uncertainty, by leveraging novel control and optimization techniques [1]. Such techniques require the knowledge of the real-time network topology, which has been traditionally lacked in many cases due to the limited visibility into the state of the system.

The inverse power flow (IPF) problem we define in this paper concerns the estimation of the nodal admittance matrix, describing the network topology, from synchronized measurements of voltage and current magnitudes and phase angles which can be obtained from phasor measurement units (PMUs) [2] or conventional supervisory control and data acquisition (SCADA) technology. The IPF problem is a host of several crucial applications, affecting real-time system operation and long-term planning, the most important of which are:

- **State Estimation** combines the knowledge of the admittance matrix with a set of known state-variables to determine the unknown ones, e.g., voltage magnitude and phase angle at unobserved buses, thereby building a real-time model of the network which enables the operators to justify technical and economical decisions and uncover potential operational problems.
- **Optimization & Control** techniques determine a sequence of operations that can transition the power system from one steady state to another one that meets certain efficiency or stability targets. These techniques typically require the knowledge of the network topology or information about the state of the system.
- **Event Detection** concerns identifying faults, line outages, and other events, such as transformer tap, capacitor, and switching operations from changes in the real-time network model.
- **Cybersecurity** is the problem of identifying the potential vulnerabilities of a power system and designing strategies to secure it from the potential cyberattacks using telemetry data along with the information about its topology.

Contributions: In this paper, we lay out the theoretical foundation for the IPF problem and explore its application in detecting faults and other critical events. Using the bus injection model (BIM), we propose efficient algorithms to identify the admittance matrix. In particular, we show that when the system has no hidden state the admittance matrix can be uniquely identified from a sequence of complex voltage and current measurements corresponding to different steady states. Should there be some hidden states in the network, we show that a reduced admittance matrix

(from Kron reduction) can be determined; we develop the following algorithms for identifying the admittance matrix of the original system: a) for radial networks, a graph-theoretical approach based on graph decomposition, maximal clique searching and composition, b) for mesh networks, an algorithm based on low rank and sparse matrix decomposition. Furthermore, a convex relaxation has been proposed to obtain a computationally efficient algorithm. Power flow simulations are performed on the IEEE 14-bus system to back up the theoretical results and evaluate their sensitivity to the measurement noise which is commonly introduced by transducers.

The paper is outlined as follows – we formulate the IPF problem and propose a solution for the case that the system is fully observable in Section II. When the system has some hidden nodes, we propose two efficient algorithms to solve the IPF problem for radial and mesh networks in Section III. We study the fault detection problem in Section IV and evaluate the identification accuracy of the proposed algorithms in Section V. We survey related work in Section VI and present directions for future work in Section VII.

II. INVERSE POWER FLOW (IPF) PROBLEM

Let \mathbb{C} denote the set of complex numbers, \mathbb{R} the set of real numbers, and \mathbb{N} the set of integers. For $A \in \mathbb{C}^{n \times n}$, $\text{Re } A$ and $\text{Im } A$ denote matrices with the real and imaginary parts of A , respectively. The transpose of a matrix A is denoted A^T and its Hermitian (complex conjugate) transpose is denoted A^H .

A power system can be modeled by a connected undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ where $\mathcal{N} := \{1, 2, \dots, N\}$ represents the set of buses, and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ represents the set of lines, each connecting two distinct buses. A bus $j \in \mathcal{N}$ can be a load bus, a generator bus, or a swing bus. Let V_j be the complex voltage at bus j and s_j be the net complex power injection (generation minus load) at that bus. We use s_j to denote both the complex number $p_j + iq_j$ and the real pair (p_j, q_j) depending on the context. For each line $(i, j) \in \mathcal{E}$ we denote its admittance by y_{ij} . The nodal admittance matrix of this system is denoted Y , which is an $N \times N$ complex-valued matrix whose off-diagonal elements are $Y_{ij} = -y_{ij}$ and diagonal elements are $Y_{ii} = -\sum_{j \neq i} Y_{ij}$, assuming that there is no shunt element. Hence, the current injection vector can be expressed as $I = YV$. The bus injection model (BIM)¹ is defined by the following power flow equations

¹We use the bus injection model in this paper; however, our approach can be generalized to other models, e.g., the DC power flow model or the Distflow model [3].

describing the Kirchhoff's law for a given time index, $k \in \{1, \dots, K\}$:

$$s_i(k) = \sum_{j \in \mathcal{N}_i} y_{ij}^H (|V_i(k)|^2 - V_i(k)V_j^H(k)), \quad \forall i \in \mathcal{N}, \quad (1)$$

where \mathcal{N}_i is the set of nodes directly connected to node i . Rewriting this formula in vector form for all time indices yields the following equation for a bus i :

$$\underbrace{\begin{bmatrix} \frac{s_i(1)}{V_i(1)} \\ \frac{s_i(2)}{V_i(2)} \\ \vdots \\ \frac{s_i(K)}{V_i(K)} \end{bmatrix}}_{I_i^{K,H}} = \underbrace{\begin{bmatrix} V_1(1) & V_2(1) & \dots & V_N(1) \\ V_1(2) & V_2(2) & \dots & V_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ V_1(K) & V_2(K) & \dots & V_N(K) \end{bmatrix}}_{V^K} \underbrace{\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{iN} \end{bmatrix}}_{Y_i}, \quad (2)$$

The IPF concerns the following mathematical problem: given voltage and power (or current) measurements of different steady-states, i.e., $V_i(k)$ and $s_i(k)$ for $k = 1, \dots, K$, whether it is possible to recover the admittance matrix Y . If so, how many steady-state measurements are required to uniquely identify Y ; otherwise, what part of Y can be identified and where additional sensors should be installed so that we can identify the entire admittance matrix.

In the following we propose different methods for solving the IPF problem when complex voltage and current measurements are available for every bus in the system. We formulate the identification problem as a constrained least square problem and convert it to an equivalent unconstrained least square problem. We note that Y has a certain structure that can be exploited when solving the IPF problem— (a) Y must be a symmetric complex matrix (i.e., $Y \in \mathbb{S}^N$) and (b) Y encodes the topology of a connected graph (or a connected tree for radial networks).

The admittance matrix Y can be obtained from solving the optimization problem below:

$$\min 0, \quad \text{s.t.:} \quad (3)$$

$$V^K Y = I^{K,H}, \quad Y \in \mathbb{S}^N$$

$$Y_{ii} = - \sum_{j \neq i} Y_{ij}, \quad \forall i$$

in which $I^{K,H} = \begin{bmatrix} I_1^{K,H} & I_2^{K,H} & \dots & I_N^{K,H} \end{bmatrix}$. This feasibility problem can be also written as a

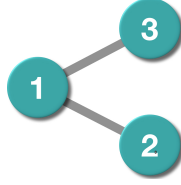


Fig. 1: An example three-node power system with two lines connecting bus 1 to buses 2 and 3.

constrained least squares problem:

$$\begin{aligned} \hat{Y}^{K,l_2} &= \arg \min_{Y \in \mathbb{C}^{N \times N}} \|V^K Y - I^{K,H}\|_F \\ \text{s.t.: } Y &\in \mathbb{S}^N, \quad Y_{ii} = - \sum_{j \neq i} Y_{ij}, \quad \forall i, \end{aligned} \quad (4)$$

Applying the *vec* operator to the objective function and constraints of the above problem, we obtain:

$$\begin{aligned} \min_{\text{vec}(Y) \in \mathbb{C}^{N^2 \times 1}} & \| (I \otimes V^K) \text{vec}(Y) - \text{vec}(I^{K,H}) \|_2 \\ \text{s.t.: } Y &\in \mathbb{S}^N, \quad Y_{ii} = - \sum_{j \neq i} Y_{ij}, \quad \forall i. \end{aligned} \quad (5)$$

Let $f : \mathbb{C}^{N \times N} \rightarrow \mathbb{C}^{(N^2-N)/2 \times 1}$ be a mapping from a symmetric complex matrix to a complex vector defined as:

$$f(Y) = \begin{bmatrix} Y_{21} & Y_{31} & \dots & Y_{N1} & Y_{32} & Y_{42} & \dots & Y_{NN-1} \end{bmatrix}^T.$$

It can be readily seen that f is a bijection for any $Y \in \mathbb{S}^N$. Based on this definition, we have $\text{vec}(Y) = Qf(Y)$, where $Q \in \mathbb{R}^{N^2 \times (N^2-N)/2}$ maps $f(Y)$ to the vectorized admittance matrix as illustrated below.

Example 1. For the network depicted in Figure 1, the Q matrix has the following form

$$\begin{bmatrix} Y_{11} \\ Y_{21} \\ Y_{31} \\ Y_{12} \\ Y_{22} \\ Y_{32} \\ Y_{13} \\ Y_{23} \\ Y_{33} \end{bmatrix} = \underbrace{\begin{bmatrix} -1 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ -1 & 0 & -1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & -1 \end{bmatrix}}_Q \begin{bmatrix} Y_{21} \\ Y_{31} \\ Y_{32} \end{bmatrix}.$$

Based on the definition of Q , the constrained l_2 optimization problem can be converted to an unconstrained l_2 optimization:

$$\min_{f(Y) \in \mathbb{C}^{(N^2-N)/2 \times 1}} \left\| \underbrace{(I \otimes V^{K,H}) Q}_{M} f(Y) - \text{vec}(I^K) \right\|_2 \quad (6)$$

The following lemma and algorithm explain how this optimization problem can be solved.

Lemma 1. *If $V^{K,H}$ has full column rank, then $I \otimes V^{K,H}$ has full column rank.*

Proof: This is easy to show using the definition of Kronecker product. ■

We define

$$\tilde{M} = \begin{bmatrix} \text{Re } M & -\text{Im } M \\ \text{Im } M & \text{Re } M \end{bmatrix}, \text{ and } \tilde{b} = \begin{bmatrix} \text{Re } \text{vec}(I^K) \\ \text{Im } \text{vec}(I^K) \end{bmatrix}.$$

The optimization problem (6) can be written as

$$\min_{\tilde{f}(Y) \in \mathbb{R}^{(N^2-N) \times 1}} \left\| \tilde{M} \tilde{f}(Y) - \tilde{b} \right\|_2, \quad (7)$$

in which $\tilde{f}(Y) \triangleq [f(Y_r)^T f(Y_i)^T]^T$, $Y_r = \text{Re } Y$, and $Y_i = \text{Im } Y$.

We compute the solution of the original optimization problem (6) from the solution of the optimization problem (7) by taking the inverse map of \tilde{f} . A sufficient condition to guarantee the exactness of the solution is that \tilde{M} has full column rank.

Proposition 1 (Exactness). *If $K \geq N$ and $V^{K,H}$ has full column rank, the optimization problem (7) has a unique solution.*

Proof: Since $Q \in \mathbb{R}^{N^2 \times (N^2 - N)/2}$ and has full column rank (this can be checked easily), there exists a matrix Q^\dagger such that $Q^\dagger Q = I$. For the Kronecker product $I \otimes V^{K,H} \in \mathbb{C}^{KN \times N^2}$, $I \otimes V^{K,H}$ has full column rank when $K \geq N$; therefore, \tilde{M} and M have full column rank given the fact that $\text{rank}(\tilde{M}) = 2\text{rank}(M)$.

Finally, we prove by contradiction that if \tilde{M} has full column rank, the solution to the optimization problem (7) is unique. Suppose there exists $\tilde{f}(Y_1)$ and $\tilde{f}(Y_2)$ ($\tilde{f}(Y_1) \neq \tilde{f}(Y_2)$) such that $\tilde{M}\tilde{f}(Y_1) = \tilde{b}$ and $\tilde{M}\tilde{f}(Y_2) = \tilde{b}$, then

$$\tilde{M} \left(\tilde{f}(Y_1) - \tilde{f}(Y_2) \right) = 0$$

which contradicts the full column rank assumption. ■

Remark 1. *Shunt elements are not considered here; nevertheless, our approach can be easily extended to the case that there are some shunt elements by changing the definition of $f(\cdot)$ and Q .*

We can add the element-wise positivity constraint $f \geq 0$ to this problem if the conductance and susceptance of each line are positive².

$$\min_{\tilde{f}(Y) \geq 0} \left\| \tilde{M}\tilde{f}(Y) - \tilde{b} \right\|_2. \quad (8)$$

The above problem is known as nonnegative least squares and can be solved using different methods, such as the active set method [4] which is described in Algorithm 1. Specifically, this algorithm takes (a) a real-valued matrix A of dimension $m \times n$, which is \tilde{M} in our case, (b) a real-valued vector y of dimension m , which is \tilde{b} , and (c) a real number t which is the tolerance used as the stopping criterion.

III. INVERSE POWER FLOW WITH HIDDEN NODES

In the previous section the IPF problem is solved for the ideal case where complex voltage and current measurements are available for all buses. However, in practice not all buses are equipped with PMUs and there might be several unobserved (hidden) nodes. For example, distribution systems typically have only a few measurement nodes installed at the substation and all other

²We know from physics that the conductance of a line is always positive, but its susceptance can be positive or negative depending on its inductive and capacitive reactance values.

Algorithm 1 Active Set Algorithm

```

1: Set  $P = \{0\}$ 
2: Set  $R = \{1, \dots, n\}$ 
3: Set  $x$  to a zero vector of dimension  $n$ 
4: Set  $w = A^T(y - Ax)$ 
5: while  $R \neq \{0\}$  and  $\max(w) > t$  do
6:   Let  $j$  be the index of  $\max(w)$  in  $w$ 
7:   Add  $j$  to  $P$  and remove  $j$  from  $R$ 
8:   Let  $A_P$  be  $A$  restricted to the variables included in  $P$ 
9:   Let  $S_P = (A_P)^\dagger y$ 
10:  while  $\min(S_P) \leq 0$  do
11:    Let  $\alpha = \min(\frac{x_i}{x_i - s_i})$  for  $i$  in  $P$  where  $s_i \leq 0$ 
12:    Set  $x$  to  $x + \alpha(s - x)$ 
13:    Move to  $R$  all indices  $j$  in  $P$  such that  $x_j = 0$ 
14:    Set  $s_P = (A_P)^\dagger y$ 
15:    Set  $s_R$  to zero
16:  end while
17:  Set  $x = s$  and  $w = A^T(y - Ax)$ 
18: end while

```

buses are hidden, whereas transmission systems are mostly covered by the measurement nodes and there might be very few hidden nodes. In any case, solving the IPF problem with hidden nodes is more challenging due to the intrinsic difficulties explained in Figure 2. Specifically, the left figure shows a power system with a single hidden node and a sparse topology. Even though there is only one hidden node, the inferred network topology from the measurements will be a full mesh as shown in the middle figure. The right figure shows the topology of the network interconnecting the observed nodes. A desired identification algorithm should return this topology, neglecting the additional links that are introduced because of the hidden node.

1) *Kron Reduction*: We make an important assumption that all buses with non-zero net current injection are equipped with measurement devices; therefore, the net power and current³ injection is always zero at a hidden node, implying that neither loads nor generators are connected to it. This assumption is necessary for identification.

Let M_1 and \bar{M}_1 represent the set of observed and hidden nodes, respectively, and h be the number of hidden nodes. We have $M_1 \cap \bar{M}_1 = \{0\}$ and $M_1 \cup \bar{M}_1 = \{1, 2, \dots, N\}$. For $i \in M_1$, the injected power s_i^k and the voltage V_i^k can be measured at different time indices k ; while,

³If the bus voltage is known, the injected current can be computed from injected power measurements $I_i(k) = \frac{s_i(k)}{V_i(k)}$ and vice versa.

for $i \in \bar{M}_1$, we have $s_i^k = I_i(k) = 0$, $\forall k$ and V_i^k is also not measured. We partition Y into four sub-matrices such that $Y_{22} \in \mathbb{C}^{h \times h}$ corresponds to the mutual admittance of the hidden nodes only:

$$\begin{bmatrix} I_1 \\ 0 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}.$$

Theorem 1. *A reduced admittance matrix $\bar{Y} = Y_{11} - Y_{12}Y_{22}^{-1}Y_{12}^T$ can be inferred from voltage and current time-series data of measured nodes if Y_{22} is invertible.*

Proof: We first solve for V_2

$$V_2(k) = -Y_{22}^{-1}Y_{21}V_1(k), \quad \forall k, \quad (9)$$

and then substitute it to the other equation:

$$I_1(k) = (Y_{11} - Y_{12}Y_{22}^{-1}Y_{12}^T) V_1(k), \quad \forall k. \quad (10)$$

We can apply a similar algorithm as we proposed in the previous section to the measured voltage and current time-series data to obtain \bar{Y} . ■

Remark 2. *The expression for the reduced admittance matrix \bar{Y} is known as Kron reduction [5]. Every node with zero current injections can be eliminated to produce a reduced network with fewer nodes and a corresponding admittance matrix $\bar{Y} \triangleq Y_{11} - Y_{12}Y_{22}^{-1}Y_{12}^T$ which has a lower dimension. This theorem poses a fundamental limit on what can be learned about the admittance matrix from data.*

Remark 3. *We can interpret eq. (9) as solving the state estimation problem, i.e., the voltage of the hidden nodes is computed from that of the observed nodes together with the whole admittance matrix Y . Hence, if we are able to infer the admittance matrix Y (i.e., Y_{11}, Y_{12}, Y_{22}) from the measurements, the state estimation and topology identification problems would be solved at the same time!*

In the above derivation, the invertibility of Y_{22} was assumed; we now show that Y_{22} is indeed invertible for connected graphs.

Lemma 2. *Consider a connected graph with h hidden states, Y_{22} is defined as the admittance matrix between the hidden nodes, then it is invertible if $h < N$.*

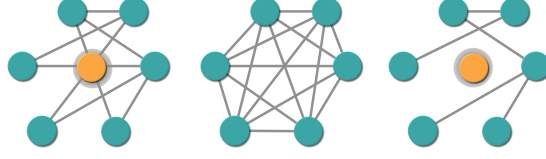


Fig. 2: **Left:** the true topology for the power system; **Middle:** the topology that we can infer from sensory data; **Right:** the topology that we would like to obtain.

Proof: Since Y_{22} is symmetric, its eigenvalues are real. Also since $-P_{22}, -Q_{22} \succ 0$ from Lemma 10 (see Appendix), we prove by contradiction that Y_{22} does not have any eigenvalue at 0. Suppose there exists a $(v + wi)$ such that $(P_{22} + Q_{22}i)(v + wi) = 0$. This is equivalent to

$$\begin{aligned} P_{22}v - Q_{22}w &= 0 \\ P_{22}w + Q_{22}v &= 0, \end{aligned} \tag{11}$$

which leads to $(P_{22} + Q_{22}P_{22}^{-1}Q_{22})w = 0$, since $-P_{22} \succ 0$ and $-P_{22}^{-1} \succ 0$, which is a contradiction. ■

We can obtain $\bar{Y} = Y_{11} - Y_{12}Y_{22}^{-1}Y_{12}^T$ from data using the least-square estimator proposed in previous section. The next question is whether we are able to estimate the original admittance matrix Y , or at least the connections between the observed nodes, i.e., Y_{11} as depicted in the right subfigure of Figure 2.

Example 2. Consider the power system shown in Figure 1, whose admittance matrix has the following structure:

$$Y(\Theta) = \begin{bmatrix} -\theta_1 - \theta_2 & \theta_1 & \theta_2 \\ \theta_1 & -\theta_1 & 0 \\ \theta_2 & 0 & -\theta_2 \end{bmatrix}.$$

If node 1 is a hidden node, which has degree 2, the reduced admittance matrix is

$$\bar{Y}(\Theta) = \frac{\theta_1\theta_2}{\theta_1 + \theta_2} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \triangleq \begin{bmatrix} -\theta_0 & \theta_0 \\ \theta_0 & -\theta_0 \end{bmatrix}.$$

A new graph with two nodes (node 2 and 3) can be constructed such that the equivalent admittance matrix is the same as the original one. In this case, we cannot do this since it violates the minimality assumption for the original graph.

For hidden nodes with degree 1, we also consider the system shown in Figure 1, if node 3 is

the hidden node, then the reduced admittance matrix is

$$\bar{Y}(\Theta) = \begin{bmatrix} -\theta_1 & \theta_1 \\ \theta_1 & -\theta_1 \end{bmatrix}.$$

Similarly, we can construct a new graph with two nodes while having the same admittance matrix.

The next definition follows directly from the above example.

Definition 1 (Equivalent systems under M_1). *Given two admittance matrices $Y^1 \in \mathbb{C}^{N_1 \times N_1}$ and $Y^2 \in \mathbb{C}^{N_2 \times N_2}$ with the same set of observed nodes M_1 , if $\bar{Y}^1 = \bar{Y}^2$, we say that these two admittance matrices are equivalent under M_1 .*

In the above definition, N_1 can be equal to N_2 . In such cases, we cannot differentiate those two systems unless we use some external information. Hence, to ensure identifiability, we make the following assumption:

Assumption 1. *Given \bar{Y} , we assume that the original admittance matrix Y corresponds to the system that has the smallest number of hidden nodes, i.e., there exists no equivalent system under M_1 : $Y^1(\cdot) : \Omega \rightarrow \mathbb{C}^m$ with $m < n$.*

2) *Sensor Placement Strategy:* To find the candidate nodes for installing the sensors, we apply Kron reduction to remove the hidden nodes with degree 2. The idea is to start with an undirected graph \mathcal{G} and a set of observed nodes. For any node that is not in this set, we remove it and add new edges (according to the rule specified in Algorithm 2) to the graph.

A new admittance matrix is then obtained for the reduced graph. We hereby define a Schur complement of a matrix $A \in \mathbb{C}^{n \times n}$

We repeat this process until all the hidden nodes are removed from the graph.

Algorithm 2 Graph Condensation Algorithm

- 1: **Input:** a graph \mathcal{G} with n nodes and a set of observed nodes $M = \{1, 2, \dots, m\}$
 - 2: **for** $v = m + 1 : n$ **do**
 - 3: $\mathcal{G} = \mathcal{G}/v$
 - 4: $\forall w, l \in \mathcal{N}_v$ (neighbors of v) add an edge between w and l to \mathcal{G} .
 - 5: **end for**
 - 6: **return** $\mathcal{G}_c = \mathcal{G}$
-

We should be able to uniquely recover the admittance matrix of the reduced graph by following the algorithm proposed in the previous section. Hence, a sufficient condition for sensor placement is that the PMUs must be installed at the hidden nodes with degree less than 3.

The power system in Example 2 has 3 nodes (2 observed and 1 hidden). From voltage and current data, we can construct a network with fewer nodes (for example 2 nodes) which has the same admittance matrix. In this case, we are not able to uncover the true topology unless additional information is available. We generalize this idea in the following proposition.

Proposition 2. *Hidden nodes with degree less than 3 are not identifiable from voltage and current measurements.*

Proof. If not, we can construct an equivalent network under measured nodes M_1 with less number of nodes as in Example 2. \square

We are therefore interested in finding whether the topology of a system with hidden nodes of degree ≥ 3 is identifiable. Once solved, it will provide a necessary condition for the sensor placement problem. In the next subsections, we propose two algorithms for identifying the original admittance matrix of radial and mesh networks.

A. Radial Networks

1) *Preliminary on graph theory:* Consider an undirected graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with $\mathcal{N} := \{1, \dots, N\}$. Two nodes j and k are *adjacent* if $j \sim k \in \mathcal{E}$. A *complete* graph is the one in which all nodes are adjacent. A subgraph of \mathcal{G} is a graph $\mathcal{F} = (\mathcal{N}', \mathcal{E}')$ with $\mathcal{N}' \subseteq \mathcal{N}$ and $\mathcal{E}' \subseteq \mathcal{E}$. A *clique* of \mathcal{G} is a complete subgraph of \mathcal{G} . A *maximal clique* of \mathcal{G} is a clique that is not a subgraph of another clique of \mathcal{G} .

For a symmetric admittance matrix $Y \in \mathbb{C}^{N \times N}$, we can define its corresponding graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with $\mathcal{N} := \{1, \dots, N\}$ and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ such that for any pair of nodes (k, j) , there exists an edge if and only if $Y[k, j] \neq 0$.

For two graphs $\mathcal{G}_1 = (\mathcal{N}, \mathcal{E}_1)$ and $\mathcal{G}_2 = (\mathcal{N}, \mathcal{E}_2)$ with the same number of nodes and the same ordering, $\mathcal{N} := \{1, \dots, N\}$ and $\mathcal{E}_1, \mathcal{E}_2 \subseteq \mathcal{N} \times \mathcal{N}$. We define $\mathcal{G}_3 = \mathcal{G}_1 / \mathcal{G}_2 = (\mathcal{N}, \mathcal{E}_3)$ such that \mathcal{E}_3 contains all the edges in \mathcal{E}_1 which do not belong to \mathcal{E}_2 and define $\mathcal{G}_4 = \mathcal{G}_1 \cup \mathcal{G}_2 = (\mathcal{N}, \mathcal{E}_4)$ such that \mathcal{E}_4 contains all the edges which belong to either \mathcal{E}_1 or \mathcal{E}_2 . We say that there is no overlapping between two graphs \mathcal{G}_1 and \mathcal{G}_2 if $\mathcal{G}_1 / \mathcal{G}_2 = \mathcal{G}_1$.

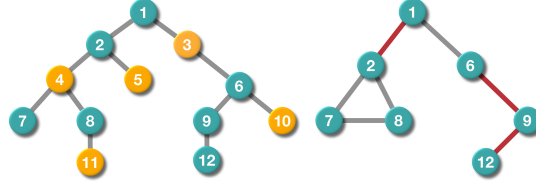


Fig. 3: An example of a 12-bus network. **Left:** The original network topology. Green circles denote the observed nodes while yellow ones denote the hidden nodes. **Right:** The topology of the condensed graph obtained from Kron reduction. The red lines denote the edges in \mathcal{G}_1 and the grey ones denote the ones in \mathcal{G}_2 .

2) *Separating graphs \mathcal{G}_1 and \mathcal{G}_2 :* Given a graph \mathcal{G} , let $\mathcal{G}_1 = (\mathcal{N}, \mathcal{E}_1)$ be the graph with observed nodes with corresponding admittance matrix $Y_{11} - \text{diag}\{Y_{11}\}$ and $\mathcal{G}_2 = (\mathcal{N}, \mathcal{E}_2)$ be the graph with new edges introduced by the Kron reduction with corresponding admittance matrix $\text{diag}\{Y_{11}\} - Y_{12}Y_{22}^{-1}Y_{21}$. From data, one can compute $\mathcal{G}_4 = \mathcal{G}_1 \cup \mathcal{G}_2$ with admittance matrix $\bar{Y} = Y_{11} - Y_{12}Y_{22}^{-1}Y_{21}$, the next question is whether \mathcal{G}_1 and \mathcal{G}_2 can be identified from \mathcal{G}_4 or equivalently $Y_{11} - \text{diag}\{Y_{11}\}$ and $\text{diag}\{Y_{11}\} - Y_{12}Y_{22}^{-1}Y_{21}$ from \bar{Y} .

We consider the reverse process from a graph-theoretical perspective.

Lemma 3. *For radial networks, let $\mathcal{G}_3 \triangleq (\mathcal{N}, \mathcal{E}_3) = \mathcal{G}_1 \cap \mathcal{G}_2$, then $\mathcal{E}_3 = \{0\}$.*

Proof: If not, there exists an edge e , without loss of generality between node 1 and node 2, that belongs to both \mathcal{G}_1 and \mathcal{G}_2 then it means there exists a direct connection between node 1 and 2 and another path from node 1 and node 2 through hidden nodes which creates a loop and therefore contradiction. ■

Lemma 4. *For any two nodes in a radial network, they either have at most one shared neighbor or have no shared neighbor if they are adjacent.*

Proof: This can be easily proved by contradiction to the tree network structure. ■

One can obtain $\mathcal{G}_4 = \mathcal{G}_1 \cup \mathcal{G}_2$ from the measurements by solving problem (4). We can make these observations:

- \mathcal{G}_1 is a subgraph of \mathcal{G} (a tree), therefore it is a tree
- \mathcal{G}_2 contains only cliques from the graph condensation algorithm
- $\mathcal{G}_3 \triangleq (\mathcal{N}, \mathcal{E}_3) = \mathcal{G}_1 \cap \mathcal{G}_2$, then $\mathcal{E}_3 = \{0\}$

Based on these three facts, we prove separability of \mathcal{G}_1 and \mathcal{G}_2 from \mathcal{G}_4 . Inspired by Example 2

which showed that we are not able to uniquely identify the admittance matrix of a system with hidden nodes that are connected to one or two observed nodes, we make the following assumptions.

Assumption 2. *All hidden nodes in the network have degree ≥ 3 .*

Theorem 2 (Separability). *Given a radial network that satisfies Assumption 2, the reduced graph \mathcal{G}_2 can be decomposed into a number of cliques $\mathcal{C}_1, \dots, \mathcal{C}_k$ (where each clique has more than two nodes) and a tree \mathcal{T} , such that $\mathcal{C}_i \cap \mathcal{C}_j = \{0\}$ for any $i, j \in \{1, 2, \dots, k\}$ and $i \neq j$, and the following statements hold*

1. $\mathcal{G}_1 = \mathcal{T}$;
2. $\mathcal{G}_2 = \cup_i \mathcal{C}_i$.

Proof: From Lemma 3, there is no overlapping between the two graphs. Also, under these assumptions, \mathcal{G}_1 is a tree and \mathcal{G}_2 is a union of cliques (≥ 3), they are totally different graphs. Third, for any \mathcal{C}_i and \mathcal{C}_j , there is no overlapping between them since there is no connection between them (or no path through hidden nodes). Combining these three properties, we finish the proof. ■

Corollary 1. *Given a radial network that satisfies Assumption 2 and if there is no direct connection between two hidden nodes, the reduced graph \mathcal{G}_2 can be decomposed to a number of cliques $\mathcal{C}_1, \dots, \mathcal{C}_k$ (with the number of nodes ≥ 3) and a tree \mathcal{T} , such that $\mathcal{C}_1 \cap \mathcal{C}_2 = \{0\}$ and the following statements hold:*

1. $\mathcal{G}_1 = \mathcal{T}$;
2. $\mathcal{G}_2 = \cup_i \mathcal{C}_i$;
3. *the number of cliques and the number of hidden states are equal, i.e., $k = h$.*

Remark 4. *Decomposing a radial network into a number of cliques and a tree allows us to obtain the true topology between the observed nodes and between the hidden nodes.*

Finding the maximum clique of a graph is termed the clique problem and is NP-complete in general. There are many algorithms for solving the clique problem, such as the Bron-Kerbosch algorithm, which we adopt here.

3) *Obtaining Y_{22} and Y_{12} :* Once \mathcal{G}_1 and \mathcal{G}_2 are separated from \mathcal{G} , we propose another algorithm to obtain Y_{11} , Y_{22} and Y_{12} .

Algorithm 3 Graph Decoupling Algorithm

```

1: Input: a condensed graph  $\mathcal{G}$ 
2: Set  $\mathcal{G}' = \mathcal{G}$ .
3: while  $\mathcal{G}'$  has a clique do
4:   Use Bron-Kerbosch algorithm to find a clique ( $\geq 3$  nodes)  $\mathcal{C}_i$  in  $\mathcal{G}'$ ,
5:   Let  $\mathcal{G}' = \mathcal{G}'/\mathcal{C}_i$ 
6: end while
7: return  $\mathcal{G}_2 = \cup_i \mathcal{C}_i$  and  $\mathcal{G}_1 = \mathcal{G}/\mathcal{G}_2$ 

```

The idea of the proposed algorithm is as follows: we first obtain the admittance matrix of \mathcal{G}_1 , denoted \tilde{Y}_{11} . From the separation results in the previous section, $\tilde{Y}_{11} = Y_{11} - \text{diag}\{Y_{11}\}$. We then decompose \mathcal{G}_2 which has an admittance matrix $\text{diag}\{Y_{11}\} - Y_{12}Y_{22}^{-1}Y_{12}^T$ to obtain Y_{22} and Y_{12} . Without loss of generality, we assume that there exists only one clique; otherwise, we can repeatedly apply the following algorithm for every clique and combine the resulting matrices together to obtain Y_{22} and Y_{12} .

More specifically, we introduce a few useful technical lemmas before presenting the main algorithm.

Lemma 5. *For any \mathcal{C}_i , an observed node can be connected to only one hidden node.*

Proof. This can be proved by contradiction. If this does not hold, there exists a loop since there is a path between hidden nodes in \mathcal{C}_i . □

For any pair of nodes k and j ($k, j \in \{1, 2, \dots, n\}$) we define $\beta[k, j]$ as the element-wise ratio between the k th and j th column of \bar{Y} . We set $\beta[k, j] = 1$ when there exists nonzero ratio $\alpha[k, j]$ such that $\bar{Y}[k, :] = \alpha[k, j] * \bar{Y}[j, :]$ except the k th and j th elements; otherwise, $\beta[k, j] = 0$.

From Lemma 5, we can prove the following lemma which provides criteria for detecting hidden nodes in the system.

Lemma 6. *For any $k \neq j$, $\beta[k, j] = 1$ if and only if node k and node j are connected to the same hidden node.*

Proof. This can be easily shown. □

We develop Algorithm 4 for recovering \hat{Y} from \bar{Y} . It is easy to see that

$$\text{diag}\{Y_{11}\} = \hat{Y}_{11} \quad (12)$$

$$Y_{11} = \tilde{Y}_{11} + \hat{Y}_{11}. \quad (13)$$

Algorithm 4 Obtain \hat{Y}

```

1: Input: a clique  $\mathcal{C}$ , its corresponding admittance matrix  $\bar{Y}$ 
2: for any pair of nodes  $(j, k)$  do
3:   Compute  $\beta[j, k]$ . If  $\beta[j, k] = 1$ , compute  $\alpha[j, k]$ 
4: end for
5: for  $j, k = 1 : h$  do
6:   if  $\beta[j, k] = 1$  then
       
$$\hat{Y}_{11}[j, j] = \bar{Y}_{11}[j, j] - \bar{Y}_{11}[j, k]/\alpha[j, k]$$

       
$$\hat{Y}_{11}[k, k] = \bar{Y}_{11}[k, k] - \bar{Y}_{11}[k, j] * \alpha[j, k]$$

   end if
7:   Compute the weight for the links connecting observed nodes to hidden nodes by
   parameterizing a matrix  $X$ .
8:   Solve  $\hat{Y}_{11} - \bar{Y} = \hat{Y}_{12}X\hat{Y}_{12}^T$  for  $X$ .
9: end for
10: Set  $\hat{Y}_{22} = X^{-1}$ 
11: Set  $\hat{Y} = \begin{bmatrix} \hat{Y}_{11} & \hat{Y}_{12} \\ \hat{Y}_{21} & \hat{Y}_{22} \end{bmatrix}$ 
12: if the graph corresponding to  $\hat{Y}_{22}$  contains a loop then
13:   Repeat the above process, treating  $\hat{Y}_{22}$  as  $\bar{Y}$ 
14:   Obtain  $\hat{Y}_{22,11}, \hat{Y}_{22,12}, \hat{Y}_{22,22}$  such that
       
$$\hat{Y}_{22} = \hat{Y}_{22,11} - \hat{Y}_{22,12}\hat{Y}_{22,22}^{-1}\hat{Y}_{22,21}.$$

15: end if
16: Set  $\hat{Y} = \begin{bmatrix} \hat{Y}_{11} & \hat{Y}_{12} & 0 \\ \hat{Y}_{21} & \hat{Y}_{22,11} & \hat{Y}_{22,12} \\ 0 & \hat{Y}_{22,21} & \hat{Y}_{22,22} \end{bmatrix}$ 
17: return  $\hat{Y}$ 

```

Finally, we apply Algorithm 4 to obtain \hat{Y} . The following example summarizes the above results.

Example 3. Given the graph shown in Figure 3 (left), if sensors are installed at nodes $\{1, 2, 6, 7, 8, 9, 12\}$, then we can use Kron reduction to obtain the reduced graph shown in Figure 3 (right). The first step is decompose the estimated \bar{Y} matrix to two graphs using Algorithm 3, one is a collection of cliques and the other one is a tree and some isolated nodes. The second step is to apply

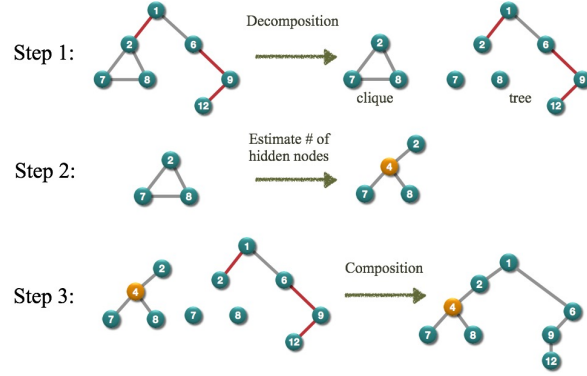


Fig. 4: A step-by-step illustration of the proposed algorithm.

Algorithm 4 to identify the original admittance matrices for all cliques. The final step is to take the union of the cliques and the graph obtained in the second step. These steps are shown in Figure 4.

B. Mesh Networks

For mesh networks, the above graph theoretical algorithm does not work since some of the properties that the algorithm relies on do not hold when there exists a loop in the network. Instead, we propose algebraic solutions to such a problem. The first step is similar to the case without hidden nodes: \bar{Y} is computed from successive current and voltage measurements. When there exist hidden nodes, inferring Y from \bar{Y} is an ill-posed problem since there often exist more variables than equality constraints. We therefore need to resort to additional prior knowledge about the original system to constrain the solution space.

Interestingly, some prior knowledge about \bar{Y} is available: first, Y_{11} corresponds to a tree and is therefore sparse; second, $Y_{12}Y_{22}^{-1}Y_{12}^T$ has low rank (it has a maximal rank equal to the number of hidden states h). We consider the following optimization problem in which \bar{Y} is decomposed into a sparse matrix A and a low-rank matrix B :

$$\begin{aligned} \begin{bmatrix} Y_{11} & Y_{12}Y_{22}^{-1}Y_{12}^T \end{bmatrix} &= \arg \min_{A,B} \|A\|_0 + \lambda \text{rank}(B) \\ \text{s.t.: } A - B &= \bar{Y}. \end{aligned} \quad (14)$$

Here $\lambda \in \mathbb{R}^+$ is a weighting parameter that balances the sparsity of Y_{11} and the number of hidden states, i.e., h . So we penalize the sparsity of Y_{11} and a low rank (a small number of hidden states) $Y_{12}Y_{22}^{-1}Y_{12}^T$ by separating these two matrices from their summation.

Without the knowledge of the true admittance matrix, it is impossible to have any identifiability property as discussed in [6]. In what follows, we propose a computationally efficient algorithm (convex relaxation) which relaxes (14). We then study how good such a relaxation is.

We can relax this problem to the following form

$$\begin{aligned} \min \quad & \|A\|_1 + \lambda \|B\|_* \\ \text{s.t.} \quad & A - B = \bar{Y}. \end{aligned} \tag{15}$$

We use l_1 optimization as a convex relaxation of l_0 optimization, and nuclear norm $\|B\|_* \triangleq \sum_i \sqrt{\lambda_i(B^H B)}$ as a convex relaxation of the rank optimization.

The problem (15) can be recast as a semidefinite program (SDP). To solve it, the first step is to use the fact that the spectral norm $\|\cdot\|$ is the dual norm of the nuclear norm $\|\cdot\|_*$ [6]:

$$\|M\|_* = \max\{\text{trace}(M^T Y) \mid \|Y\| \leq 1\}.$$

Further, the spectral norm admits a simple semidefinite characterization:

$$\|Y\| = \min_t \quad t \quad \text{s.t.} \quad \begin{pmatrix} tI_n & Y \\ Y^T & tI_n \end{pmatrix} \succeq 0.$$

From duality, we can obtain the following SDP characterization of the nuclear norm and problem (15) can be rewritten as

$$\begin{aligned} \min_{A, B, W_1, W_2} \quad & \gamma \|A\|_1 + \frac{1}{2}(\text{trace}(W_1) + \text{trace}(W_2)) \\ \text{s.t.} \quad & \begin{pmatrix} W_1 & B \\ B^T & W_2 \end{pmatrix} \succeq 0 \\ & A - B = \bar{Y}. \end{aligned} \tag{16}$$

Lemma 7. *Given a complex matrix M and consider a constructed \tilde{M}*

$$\tilde{M} = \begin{bmatrix} \text{Re } M & -\text{Im } M \\ \text{Im } M & \text{Re } M \end{bmatrix}, \tag{17}$$

we have the following properties:

- $\text{rank}(\tilde{M}) = 2\text{rank}(M)$;
- $\|\tilde{M}\|_* = 2\|M\|_*$.

Proof: Easy to show. ■

Based on the above lemma, the optimization problem (16) can be converted to the following convex optimization:

$$\begin{aligned}
 \min_{A, B, \tilde{W}_1, \tilde{W}_2, Z} \quad & \gamma \mathbf{1}_n^T Z \mathbf{1}_n + \frac{1}{2} (\text{trace}(\tilde{W}_1) + \text{trace}(\tilde{W}_2)) \\
 \text{s.t.:} \quad & \begin{pmatrix} \tilde{W}_1 & \tilde{M} \\ \tilde{M}^T & \tilde{W}_2 \end{pmatrix} \succeq 0 \\
 & -Z[i, j] \leq A[i, j] \leq Z[i, j], \quad \forall (i, j) \\
 & A - B = \begin{bmatrix} \text{Re } \bar{Y} & -\text{Im } \bar{Y} \\ \text{Im } \bar{Y} & \text{Re } \bar{Y} \end{bmatrix}
 \end{aligned} \tag{18}$$

A, B can be decomposed to the form expressed in (17) for any symmetric matrix M .

Remark 5. The last constraint on A, B can be easily turned to linear constraints on their coefficients.

Once we obtain A, B from this optimization, we can obtain the complex matrices Y_{11} and $Y_{12}Y_{22}^{-1}Y_{21}$.

IV. CASE STUDY: EVENT DETECTION

Transformer tap, capacitor, and switching operations as well as some faults can alter the admittance between a few pairs of buses and may even change the network topology. In any case, the admittance matrix reflects these changes, suggesting that these events should be detectable from phasor measurements. Building on the IPF framework, in this section we design an efficient online algorithm for detecting and locating such events in a power system with hidden nodes. The proposed algorithm requires only a small amount of data, has a low false alarm rate, and focuses on detecting *persistent changes*, i.e., changes that affect the admittance matrix for several cycles, as in [7].

Given an affine parameterization of the admittance matrix $Y^{\delta(t)}(\cdot) : \Omega_1 \times \mathbb{Z} \rightarrow \mathbb{R}^n$, where

$$\delta(t) = \begin{cases} i & t < T \\ j & t \geq T \end{cases}$$

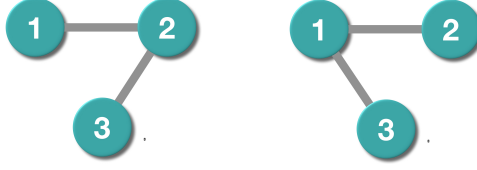


Fig. 5: Topology switches from left to the right one.

is the discrete mode, our goal is to identify the time that an event occurs, and estimate how the admittance matrix has changed from voltage and current measurements; this indicates the approximate location of the event. From Theorem 1, we can define \bar{Y}^0 and \bar{Y}^1 for the system before and after the event, respectively.

Example 4. Consider two plausible topologies of a power system illustrated in Figure 5. It is easy to verify that their reduced admittance matrices are indistinguishable if bus 3 is a hidden node. Specifically,

$$Y^0 = \begin{bmatrix} -Y_{12} & Y_{12} & 0 \\ Y_{12} & -Y_{12} - Y_{23} & Y_{23} \\ 0 & Y_{23} & -Y_{23} \end{bmatrix},$$

$$Y^1 = \begin{bmatrix} -Y_{12} - Y_{13} & Y_{12} & Y_{13} \\ Y_{12} & -Y_{12} & 0 \\ Y_{13} & 0 & -Y_{23} \end{bmatrix},$$

and the reduced admittance matrices are

$$\bar{Y}^0 = \bar{Y}^1 = \begin{bmatrix} -Y_{12} & Y_{12} \\ Y_{12} & -Y_{12} \end{bmatrix}.$$

implying that we can not detect the event that caused this change from phasor measurements at buses 1 and 2.

To ensure that the event detection problem is feasible we need to make an additional assumption:

Assumption 3. Given a power system with a set of measured nodes, we have $\bar{Y}^1 \neq \bar{Y}^0$ where \bar{Y}^0 and \bar{Y}^1 are the reduced admittance matrices of the system before and after T .

Under this assumption, the event detection and location problems can be solved even when there are some hidden states. We also exploit the intrinsic sparsity of the difference between the admittance matrices of the system before and after the change, which is due to the fact that in reality only a few elements of Y could change during an event, e.g., closing a switch and opening another one will only change two elements of Y .

A. Detecting Occurrence of Events

To detect a change in the admittance matrix, we estimate the injected current vector at time k using the known admittance matrix, \bar{Y}^0 , and the measured voltage vector at time k . We then compare the estimated injected current vector with the measured values at k to calculate the prediction error which is defined as:

$$e(k) = I(k) - \hat{I}(k) = I(k) - \bar{Y}^0 V(k), \quad (19)$$

We note that the series $e(\cdot)$ is expected to be a white noise if the admittance matrix has not changed; this can be verified using the turning point test. When the prediction error $\|e(k)\|$ becomes larger than a predefined threshold γ , we assert that the admittance matrix has changed and the reduced admittance matrix must be recomputed as described next.

B. Updating the Admittance Matrix

Upon detection of an event, we recompute the reduced admittance matrix to locate the potential fault. This can be formulated as the following optimization problem:

$$\min \|\bar{Y}^1 - \bar{Y}^0\|_0, \quad \text{s.t.: } I_1 = \bar{Y}^1 V_1, \quad (20)$$

in which $I_1 = [I(T) \ I(T+1) \ \dots \ I(T+K)]$, $V_1 = [V(T) \ V(T+1) \ \dots \ V(T+K)]$, and T is the time that the event is detected. This optimization assumes a sparse difference matrix between \bar{Y}^1 and \bar{Y}^0 as it is rarely the case that multiple events occur right at the same time. Next, we propose a relaxation of this problem. Let $\Delta Y \triangleq \bar{Y}^1 - \bar{Y}^0$. The problem can be rewritten as

$$\min \|\Delta Y\|_0, \quad \text{s.t.: } I_1 - \bar{Y}^0 V_1 = \Delta Y V_1. \quad (21)$$

Then let

$$\begin{aligned} \hat{\Delta Y}^{K,l_0} &= \arg \min_{\Delta Y \in \mathbb{C}^{N \times N}} \|\Delta Y\|_0 \quad \text{s.t.:} \\ V^{K,H} \Delta Y &= I^K - V^{K,H} \bar{Y}^0, \\ \Delta Y &\in \mathbb{S}^N, \Delta Y_{ii} = - \sum_{j \neq i} \Delta Y_{ij}, \quad \forall i. \end{aligned} \quad (22)$$

Intuitively, the optimizer of this problem is the sparsest admittance matrix difference ΔY that satisfies the same constraints as Y . A graph theoretic interpretation of this is that we are searching for a graph that has the smallest number of edges. The next proposition gives the necessary and sufficient condition for such a sparse difference matrix to converge to the true solution.

Definition 2. *The spark of a given matrix Φ , denoted $\text{Spark}(\Phi)$, is the smallest number of columns of Φ that are linearly dependent.*

Proposition 3 (From [8]). *For any vector $\mathbf{y} \in \mathbb{R}^M$, there exists one unique signal \mathbf{w} , such that $\mathbf{y} = \Phi \mathbf{w}$ with $\|\mathbf{w}\|_0 = S$ if and only if $\text{Spark}(\Phi) > 2S$.*

Remark 6. *By checking spark of a dictionary matrix Φ , we can guarantee that the solution of the l_0 optimization is the true solution. The spark of the dictionary depends on the number of samples, K , in an inexplicit way.*

We have shown so far that by checking the spark of the constructed dictionary matrix, we can examine the gap between the solution of the l_0 optimization and the true solution. Next, we propose a convex relaxation of (22), which is well-known for being computationally expensive, and specify the condition under which the solution to the l_0 optimization is equal to that of the l_1 optimization.

Following similar derivations in the l_2 optimization, we can convert the optimization problem (22) to

$$\begin{aligned} &\min_{f(\Delta Y) \in \mathbb{C}^{(N^2-N)/2 \times 1}} \|f(\Delta Y)\|_0 \\ \text{s.t.: } &\underbrace{(I \otimes V^{K,H})}_M Q f(\Delta Y) = \text{vec}(I^K). \end{aligned} \quad (23)$$

Similarly, we can define \tilde{M} , \tilde{f} and \tilde{b} , and turn the above optimization problem into

$$\min_{\tilde{f}(\Delta Y) \geq 0} \|\tilde{f}(\Delta Y)\|_0 \quad \text{s.t.: } \tilde{M} \tilde{f}(\Delta Y) = \tilde{b}, \quad (24)$$

which can be again relaxed to

$$\min_{\tilde{f}(\Delta Y) \geq 0} \|\tilde{f}(\Delta Y)\|_1, \quad \text{s.t.: } \tilde{M}\tilde{f}(\Delta Y) = \tilde{b}. \quad (25)$$

To speed up the algorithm, we can distribute the computation to every node. Specifically, node i solves the following optimization

$$\begin{aligned} \min_{\Delta Y_i \in \mathbb{C}^{N \times 1}} \|\Delta Y_i\|_0 \\ \text{s.t.: } V^{K,H}(\Delta Y_i + Y_i^0) = I_i^K, \quad \Delta Y_{ii} = - \sum_{j \neq i} \Delta Y_{ij}. \end{aligned} \quad (26)$$

In the distributed version, every node solves for its own connections using the centralized information. Similarly, we can relax the optimization and analyze the exactness of the relaxed solution.

C. Theoretical Guarantee

We now examine uniqueness of the solution obtained from l_1 optimization $\{w : Aw = b, w \geq 0, \|w\|_1 \leq \|x\|_1\} = \{x\}$.

Definition 3. (*Signum Function*): When applied to complex number $re^{i\theta}$ ($r \neq 0$), the $\text{sgn}(\cdot)$ function returns

$$\text{sgn}(re^{i\theta}) = e^{i\theta}, \text{ if } r \neq 0 \text{ and } \text{sgn}(0) = 0.$$

Theorem 3. [9] Suppose Φ is linearly independent and the sparsest representation of a complex vector is $\Phi_{opt}b_{opt}$. If the following condition holds

$$|\langle \Phi_{opt}^\dagger(\text{sgn}(b_{opt})), \varphi_w \rangle| < 1, \quad \text{for every } w \notin \Gamma_{opt}, \quad (27)$$

the unique solution of the ℓ_1 problem is b'_{opt} in which $(\cdot)'$ extends a vector to an N -dimensional one whose nonzero entries lie at coordinates index.

Based on the above derivations, we can prove the following theorem for the exactness of the solution using l_1 optimization.

Theorem 4 (Exactness). Consider the solution \hat{y}^{K,l_1} to problem (25), we have

$$\hat{Y}^{K,l_1} = f^{-1}(\hat{y}^{K,l_1}). \quad (28)$$

Algorithm 5 Event Detection & Identification Algorithm

- 1: **Input:** voltage and current data
- 2: **for** $i = 1, \dots, N$ **do**
- 3: Initialize $u_i^0 = 1, \forall j$
- 4: **for** $K = 1, \dots, k$ **do**
- 5: Solve the ℓ_1 -minimization problem

$$Y_i^{K,l_1} = \arg \min_{Y_i} \|V^{K,H}(Y_i + \bar{Y}_i^0) - I_i^K\|_2^2 + \lambda \|Y_i\|_1,$$

$$\text{s.t. : } Y_{ii} = - \sum_{j \neq i} Y_{ij};$$

- 6: **if** (27) holds and $\|Y_i^{K,l_1} - Y_i^{K-1,l_1}\|_1 \leq \epsilon$ **then**
 - 7: $\Delta Y_i = Y_i^{K,l_1}$ and **break**;
 - 8: **end if**
 - 9: **end for**
 - 10: **end for**
 - 11: **return** $\Delta Y = [\Delta Y_1 \ \Delta Y_2 \ \dots \ \Delta Y_N]$.
-

Furthermore, it recovers the true admittance matrix, i.e., $\hat{Y}^{K,l_1} = Y$ if the following conditions hold

- i) the Spark of the normalized matrix $\text{Spark}(\tilde{\Phi}) > 2S$, in which S is the sparsity of the true solution;
- ii) $|\langle \Phi_{opt}^\dagger(\text{sgn}(b^{opt})), \varphi_w \rangle| < 1$, for every $w \notin \Gamma_{opt}$.

Proof: This follows immediately from the above derivations. ■

The complex lasso problem can be solved using a greedy algorithm for example.

Remark 7. After solving the relaxed problem (25), we should check if the solution satisfies the conditions of Theorem 3. Algorithm 5 describes these steps.

V. SIMULATIONS

In this section we implement the proposed algorithms in MATLAB and evaluate their identification accuracy by performing simulations in MATPOWER [10]. The optimization problems are solved using the CVX toolbox [11]. We run power flow analysis on the IEEE 14-bus test system (Figure 6), representing a portion of a power system in the Midwestern U.S. which has 14 buses, 11 aggregated loads, and 5 generators, 3 of which are synchronous compensators used for reactive power support [12].

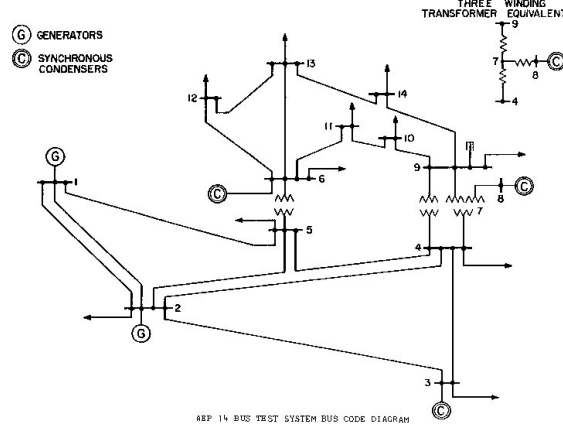


Fig. 6: The IEEE 14-bus test system [12].

We assume that PMUs are installed at selected buses and precisely measure the voltage and current values that we obtain from power flow calculations, unless stated otherwise. For each scenario, we run 100 steady state simulations, each pertaining to a time slot, to determine the voltage and current magnitude and phase angle of every bus, while varying the real and reactive power demand of the loads across the time slots. Specifically, for a given time slot, the real and reactive power consumption of a constant PQ load are computed by multiplying a scaling factor drawn from a uniform distribution over the interval $[0.8, 1.2]$ by the real and reactive power consumption data provided in [12]. We obtain the admittance matrix of this system using a built-in function of the power flow simulator. It turns out that the absolute values of nonzero complex elements of the admittance matrix are between 1.86 and 40.06, reminding the readers that a complex number's absolute value is its distance from zero in the complex plane.

We first consider the scenario that every bus is equipped with a PMU. Assuming that the self admittance of bus 7, i.e., the transformer bus, is known, Figure 7 shows the identification error, defined as $|Y - \hat{Y}|$, and the vertical color bar indicates the mapping of data values into colors. Hence, the color of a cell located at row i and column j represents the value of $|Y(i, j) - \hat{Y}(i, j)|$. It can be seen that the identification error using measurements of 15 time slots ($K = 15$) is quite small compared to the absolute value of the elements of Y , and this error does not vary much if we use more observations. We next analyze the sensitivity of the proposed algorithm to the measurement error which is typically introduced by the transducers. To this end, white Gaussian noise with a signal-to-noise ratio of 125 is added to both complex voltage and current

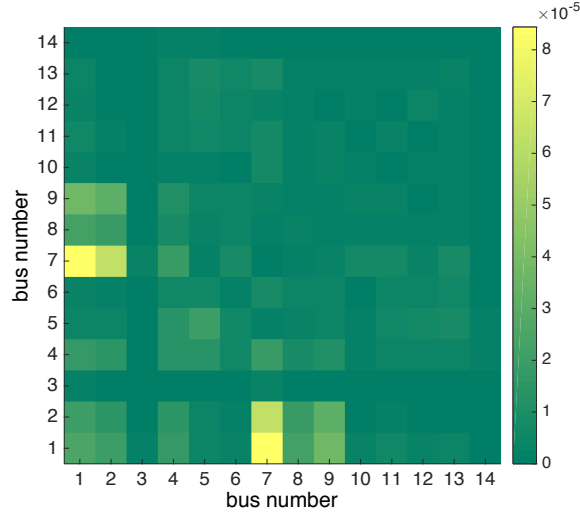


Fig. 7: The identification error when there is no hidden state.

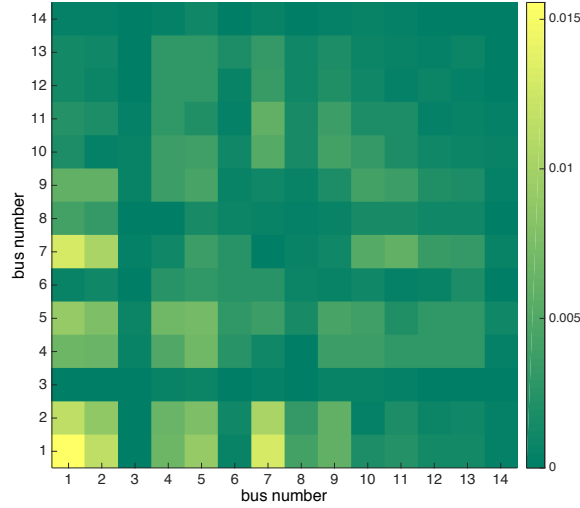


Fig. 8: The identification error when white Gaussian noise is added to both complex voltage and current measurements.

measurements. The signal-to-noise ratio is chosen such that the measurement accuracy lies within the reported range for existing PMU technology. Figure 8 shows the absolute identification error for this case. Similar to the previous case, the errors are sufficiently small. In general, we observe that the identification error increases as we decrease the signal-to-noise ratio and it becomes really large when the signal-to-noise ratio drops below 100; at this point we say that the Y matrix cannot be identified from data.

We now consider the scenario that bus 7, which has net zero injection, is a hidden node.

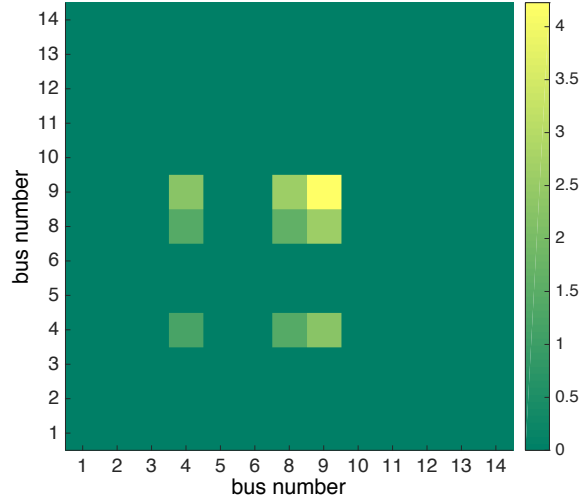


Fig. 9: The identification error for the case that bus 7 is a hidden node.

Figure 9 shows the Y_{11} identification error, which is defined as $|Y_{11} - \hat{Y}_{11}|$. It can be seen that in this case the error is relatively large compared to the absolute value of the elements of Y . This indicates that the solution of the convex relaxation is not feasible for the original l_0 optimization problem. This is a fundamental problem for identification of the admittance matrix of mesh networks with hidden states, which we plan to address in future work.

We finally consider the event detection problem. Suppose the line connecting buses 2 and 4 disconnects in time slot 50, changing the admittance between these two buses to zero. We observe that the event detection algorithm can immediately detect the occurrence of this event and its estimate of the new admittance matrix converges to the true admittance matrix using measurements of the next 6 time slots ($K = 6$), assuming that the old admittance matrix had been estimated before the event. Given the frequency at which PMUs sample voltage and current, it is possible to recover Y in a few hundred milliseconds after the event, enabling the operators to detect the event, identify its location, and take remedial actions in near real-time. Figure 10 shows the identification error, which is defined as the absolute value of element-wise error between ΔY and $\Delta \hat{Y}$ in this case.

VI. RELATED WORK

The availability of high-precision, high-sample-rate measurements of transmission and distribution networks in recent years has given impetus to research on topology identification, state estimation, and fault detection. In this section we discuss related work in these areas:

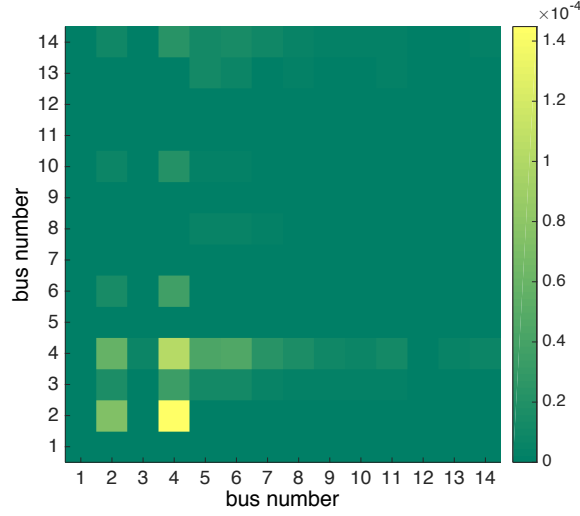


Fig. 10: The identification error using only 6 observations after the occurrence of the events.

Topology identification – The topology identification problem has been studied extensively in transmission networks [13], [14], [15] as well as distribution grids [16], [17], [18] using single-phase a.c. and d.c. power flow models. For example, the topology identification problem is cast as a sparse subspace learning problem in [13] and an efficient algorithm is proposed to estimate the admittance matrix of the underlying power system from the measured power injection of different buses. In [18], the topology of an urban (mesh) distribution network is inferred from voltage magnitude, and real/reactive power measurements carried out by smart meters at the end-nodes. A graphical model is built to describe the probabilistic relationships between different voltage measurements using lasso. In [16] a graphical learning based approach is developed to estimate the radial grid topology from nodal voltage measurements. The learning algorithm is based on conditional independence tests for continuous variables over chordal graphs. An efficient algorithm for topology identification of a power system is also proposed in [15] drawing on ideas from compressive sensing and graph theory. The authors assume that power and phase angle measurements are available for all nodes. To our knowledge, no related work addresses the topology identification problem when there is some hidden states, and this paper is the first one that solves such an identification problem with theoretical guarantees.

Event detection – Designing algorithms that can quickly detect, identify, and isolate high impedance faults using PMU data is of great importance [19]. The state-of-the-art approach is to detect a fault or an outage using either an online algorithm, which relies on a relatively small

number of observations [20], [21], [22], [7], [25], [24], or an offline algorithm [26], [23]. Our approach is more general than these approaches as it is capable of identifying *any* type of power system events that induce a change in the admittance matrix of the underlying network in quasi real-time.

Network tomography – The topology identification problem with hidden states is similar to electrical impedance tomography, in which the conductivity of a part of the body is inferred from simultaneous measurements of currents and voltages at the boundary (surface electrode measurements), as both problems concern inferring Y from \bar{Y} . It is known that the tomography problem is feasible only in highly symmetric networks [27], [28].

VII. CONCLUSIONS

This paper lays out the foundation of the inverse power flow problem which concerns inferring the admittance matrix of a power system from synchronized measurements of voltage and current obtained from a subset of its buses. The algorithms proposed in this work are efficient, robust to noise, and can jointly address state estimation and topology identification problems, if certain conditions are met. Additionally, they can be applied to detect and locate the events that induce a change in the admittance matrix, such as switching and transformer tap operations, line outages, etc, using only a small number of successive measurements. This enables the system operators to identify such events in quasi real-time and take prompt remedial actions, if necessary. These findings are supported by extensive power flow simulations performed on a standard test system.

The plausibility of our results underlines that much value can be extracted from synchrophasor data, especially in distribution systems which are not typically monitored beyond the substation. In future work, we plan to extend our framework to three phase power flow models, which take the mutual coupling between phases into account, develop efficient algorithms for identifying the admittance matrix of radial distribution systems with few measurement nodes, and analyze the sensitivity of the identification results to non-stationary measurement errors. We also intend to further examine the IPF problem with hidden states when the underlying network is a mesh and propose other efficient identification algorithms.

REFERENCES

- [1] O. Ardakanian, S. Keshav, and C. Rosenberg, *Integration of Renewable Generation and Elastic Loads into Distribution Grids*, ser. SpringerBriefs in Electrical and Computer Engineering. Springer International Publishing, 2016.

- [2] NASPI, “North American Synchrophasor Initiative,” <https://www.naspi.org>.
- [3] M. Baran and F. F. Wu, “Optimal Sizing of Capacitors Placed on a Radial Distribution System,” *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 735–743, Jan 1989.
- [4] S. Wright and J. Nocedal, “Numerical optimization,” *Springer Science*, vol. 35, pp. 67–68, 1999.
- [5] F. Dorfler and F. Bullo, “Kron Reduction of Graphs With Applications to Electrical Networks,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, Jan 2013.
- [6] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky, “Rank-Sparsity Incoherence for Matrix Decomposition,” *SIAM Journal on Optimization*, vol. 21, no. 2, pp. 572–596, 2011.
- [7] Y. C. Chen, T. Banerjee, A. D. Dominguez-Garcia, and V. V. Veeravalli, “Quickest Line Outage Detection and Identification,” *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 749–758, Jan 2016.
- [8] D. L. Donoho and M. Elad, “Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization,” *Proceedings of the National Academy of Sciences*, vol. 100, no. 5, pp. 2197–2202, 2003.
- [9] J. A. Tropp, “Recovery of short, complex linear combinations via ℓ_1 minimization,” *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1568–1570, April 2005.
- [10] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, “MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education,” *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb 2011.
- [11] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” <http://cvxr.com/cvx>, Mar. 2014.
- [12] IEEE, “14 Bus Test Case,” https://www.ee.washington.edu/research/pstca/pf14/pg_tca14bus.htm.
- [13] X. Li, H. V. Poor, and A. Scaglione, “Blind topology identification for power systems,” in *Smart Grid Communications, IEEE International Conference on*, Oct 2013, pp. 91–96.
- [14] V. Kekatos, G. B. Giannakis, and R. Baldick, “Grid topology identification using electricity prices,” in *IEEE PES General Meeting | Conference Exposition*, July 2014, pp. 1–5.
- [15] M. Babakmehr, M. G. Simoes, M. B. Wakin, and F. Harirchi, “Compressive Sensing-Based Topology Identification for Smart Grids,” *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 532–543, April 2016.
- [16] D. Deka, S. Backhaus, and M. Chertkovy, “Estimating distribution grid topologies: A graphical learning based approach,” *arXiv preprint arXiv:1602.08509*, 2016.
- [17] Y. Liao, Y. Weng, M. Wu, and R. Rajagopal, “Distribution grid topology reconstruction: An information theoretic approach,” in *North American Power Symposium*, Oct 2015, pp. 1–6.
- [18] Y. Liao, Y. Weng, and R. Rajagopal, “Urban Distribution Grid Topology Reconstruction via Lasso,” in *IEEE PES General Meeting*, July 2016, pp. x–x.
- [19] A. von Meier, D. Culler, A. McEachern, and R. Arghandeh, “Micro-synchrophasors for distribution systems,” in *IEEE PES Innovative Smart Grid Technologies Conference*, Feb 2014, pp. 1–5.
- [20] W. Pan, Y. Yuan, H. Sandberg, J. Goncalves, and G.-B. Stan, “Online fault diagnosis for nonlinear power systems,” *Automatica*, vol. 55, pp. 27–36, 2015.
- [21] L. Xie, Y. Chen, and P. R. Kumar, “Dimensionality Reduction of Synchrophasor Data for Early Event Detection: Linearized Analysis,” *IEEE Transactions on Power Systems*, vol. 29, no. 6, pp. 2784–2794, Nov 2014.
- [22] Y. Sharon, A. M. Annaswamy, A. L. Motto, and A. Chakraborty, “Topology identification in distribution network with limited measurements,” in *IEEE PES Innovative Smart Grid Technologies*, Jan 2012, pp. 1–6.
- [23] Q. Huang, L. Shao, and N. Li, “Dynamic Detection of Transmission Line Outages using Hidden Markov Models,” in *American Control Conference (ACC)*. IEEE, 2015, pp. 5050–5055.

- [24] Y. Yuan, Y. Nakahira, and C. Tomlin, "On identification of parameterized switched linear systems," in *Chinese Control Conference*, July 2016, pp. x–x.
- [25] G. Cavararo, R. Arghandeh, K. Poolla, and A. von Meier, "Data-driven approach for distribution network topology detection," in *IEEE PES General Meeting*, July 2015, pp. 1–5.
- [26] G. Cavararo, "Modeling, Control and Identification of a Smart Grid," Ph.D. dissertation, University of Padova, 2015.
- [27] L. Borcea, V. Druskin, F. G. Vasquez, and A. Mamonov, "Resistor network approaches to electrical impedance tomography," *Inverse Problems and Applications: Inside Out II, Math. Sci. Res. Inst. Publ.*, vol. 60, pp. 55–118, 2011.
- [28] E. B. Curtis, D. Ingerman, and J. A. Morrow, "Circular planar graphs and resistor networks," *Linear algebra and its applications*, vol. 283, no. 1, pp. 115–150, 1998.

APPENDIX

Lemma 8 (Gershgorin Theorem). *For $i \in \{1, \dots, n\}$, let $R_i = \sum_{j \neq i} |A[i, j]|$ and $D(A[i, i], R_i)$ be a closed disc centered at $A[i, i]$ with radius R_i , then every eigenvalue of a complex matrix A lies within at least one of the Gershgorin discs $D(A[i, i], R_i)$.* \square

Lemma 9 (Rayleigh-Ritz Theorem). *Let $A \in \mathbb{S}^h$ be a symmetric matrix with eigenvalues $\{\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_h\}$, then*

$$\lambda_k = \max\{\min\{R_A(x) \mid x \in U \text{ and } x \neq 0\} \mid \dim(U) = k\},$$

the Rayleigh-Ritz quotient $R_A : \mathbb{C}^h \rightarrow \mathbb{R}$ is defined by $R_A(x) = \frac{\langle Ax, x \rangle}{\langle x, x \rangle}$. \square

Based on these two lemmas, we can make the observation that $Y_{22} \triangleq P_{22} + iQ_{22}$, in which P_{22} and Q_{22} have the following properties from physics:

1. any non-diagonal element in P_{22} and Q_{22} is a non-negative real value due to the fact that conductance and susceptance of each line are positive;
2. $P_{22}[i, i] \geq -\sum_j P_{22}[i, j]$ and $Q_{22}[i, i] \geq -\sum_j Q_{22}[i, j]$ for any i , where equality holds when shunt elements do not exist.

Lemma 10. *Consider a connected network with h hidden nodes, if $h < N$, then $-P_{22}, -Q_{22} \succ 0$.*

Proof: First, from Gershgorin Theorem, all eigenvalues of P_{22} and Q_{22} lie in the LHP including the origin.

Second, using Rayleigh-Ritz theorem, for any vector $x \in \mathbb{R}^h$ and $\langle x, x \rangle = 1$, it has

$$\begin{aligned}
 R_{P_{22}} &= x^T P_{22} x \\
 &= \sum_{i \sim j, i, j \in M_1, i < j} P_{22}[i, j] (x_i - x_j)^2 \\
 &\quad + \sum_{i \sim k, i \in M_1, k \in \bar{M}_1} P_{22}[i, k] x_i^2.
 \end{aligned}$$

Next, we show that $R_{P_{22}} > 0$, if not, since it is a sum of squares, then $R_{P_{22}} = 0$ which leads to

1. $x_i = 0$ for any hidden node i connected to a measured node;
2. $x_i = x_j$ for $i - j$ is a connection between hidden nodes.

Since the underlying network is a connected graph, there exists a connection between measured nodes and hidden nodes. Moreover, the topology between hidden nodes is a subgraph of a tree, and therefore, it is a connected tree. From Property 2, we can see $x_j = 0$ for any $j \in \bar{M}_1$, which contradicts the definition of Rayleigh-Ritz quotient. Similarly, we can show this for Q_{22} , which concludes the proof. ■